Secularism & Nonreligion

RESEARCH ARTICLE

Finding the Missing Atheists

Ryan Burge and Hannah Smothers

Anyone who has ever written and distributed a survey knows that the process can oftentimes be drawn out and incredibly tedious. It is almost inevitable that the survey author will realize that they forgot to ask about a specific topic, did not offer enough response options, or did not fully investigate a possible question ordering effect. However, oftentimes these oversights can be seen as opportunities. In the 2010 wave of the Cooperative Congressional Election Study, when individuals were asked about their present religion, the response options did not include a choice for atheists, which is present in all other waves of the CCES. If an atheist was taking the questionnaire in 2010 and did not see the optimal choice that described their religious affiliation, what was their backup option? Using both descriptive analysis and machine learning techniques, we try to determine where these misplaced atheists went in the 2010 CCES. In general, we found that the vast majority chose one of the other religiously unaffiliated options: agnostics or nothing in particular, but a significant minority chose another religious tradition. We believe that these results help illuminate how atheists think about their religious affiliation and give researchers more insight into the religious landscape of the United States.

Introduction

Anyone who has ever written and distributed a survey knows that the process can oftentimes be drawn out and incredibly tedious. When the instrument finally goes live and responses begin to pour in, there is a sense of excitement which is often accompanied by a feeling of dread. It is almost inevitable that the survey author will realize that they forgot to ask about a specific topic, did not offer enough response options, or did not fully investigate a possible question ordering effect. Unfortunately, most of the time it is too late to correct these errors and researchers are left with trying to make sense of a survey that they consider to be imperfect or incomplete.

However, oftentimes these oversights can be seen as opportunities. Such is the case with the 2010 Cooperative Congressional Election Study. This survey is one of the most important datasets for students of American politics and religion because of its sheer size and scope. However, in the 2010 wave, when individuals were asked about their present religion, the response options did not include a choice for atheists, which is present in all other waves of the CCES. If an atheist was taking the questionnaire in 2010 and did not see the optimal choice that described their religious affiliation, what was their backup option? And is it possible to try to recreate the atheist sample? Using both descriptive analysis and machine learning techniques, we try to determine where these misplaced atheists went in the 2010 CCES. In general, we found that the vast majority chose one of the other religiously

unaffiliated options: agnostics or nothing in particular, but a significant minority chose another religious tradition. We believe that these results help illuminate how atheists think about their religious affiliation and give researchers more insight into the religious landscape of the United States.

Literature Review
Shift From Religion and Measuring "Nones"

One of the most important recent developments in religious demography has been a shift away from religion (Hout and Fischer 2002). People are identifying less with traditional religious sects which in turn suggests a general trend towards non-religious affiliation (Swatos and Christiano 1999). This has presented itself with an opportunity to find new ways to measure the religiously unaffiliated or religious "nones." This has not been without its challenges. Finding new ways of measuring the religiously unaffiliated has been difficult mainly because the standard ways of asking about religion simply do not apply to the religiously unaffiliated, making it difficult for them to answer these questions (Cragun 2019). Scholars have become especially focused on how we should conceptualize religious "nones" in recent empirical work. (Baker and Smith 2015; Zuckerman et al. 2016). This suggests that these groups require more careful analysis, and a rethinking of measurement approaches to accommodate this shift in religious affiliation.

Measurement Difficulties

One of the obstacles that all social scientists struggle with is to generate consistent and accurate measurements of social phenomena. Religious affiliation represents a

Eastern Illinois University, US

Corresponding author: Ryan Burge (ryanburge@gmail.com)

particularly difficult methodological puzzle. Issues like respondents overstating their commitment to a religious affiliation, or not fully understanding their current religious tradition can, oftentimes, create small, but notable differences in results (Caplow 1998; Kohut et al. 2001; Smith and Kim 2007). In general, survey questions are oftentimes biased against those without a religious affiliation. This is the case because these surveys only look at how religious people are, not how religious they are not (Hall, Meador, and Koenig 2008; Hall, Meador, and Koenig 2009; Hwang, Hammer, and Cragun 2011). Previously conducted research attempted to solve these measurement problems through various types of survey questions. Smith and Kim are an example of this, where they explore the nuances of Protestant Christians (Smith and Kim 2005). They note that there is not a single best way to measure Protestant affiliation, but this is further complicated by the ever-increasing plethora of world religions. This is less true if the questions are open-ended, although it leaves analysts to make judgment calls about how to sort many respondents who do not fit easily into one category (Smith 1991).

Survey administrators also run into issues of respondent's choosing answers that they believe will please the survey administrator through social desirability bias (Nederhof and Zwier 1983). This is true with any survey, across fields, that have questions regarding morality or self-assessment (Karp and Borckington 2005; McGuire 1968; Nederhof 1985; Nederhof and Zwier 1983; Streb et al. 2007). Topics like religion, sexual habits, and drug use lend themselves to this type of bias. Overstating their church attendance is a prime example of this (Hout and Greely 1998; Presser and Stinson 1998; Woodberry 1998). It is especially evident by in Hout and Greely's study in which the number of respondents saying that they attended church weekly did not correspond with the amount of people in the pews on Sundays in Ashtabula County, Ohio (Hout and Greely 1987). Other research found that the general discrepancy was somewhere between 16 and 20 percentage points (Hadaway et al. 1993, 1998). Attempts at measuring the religious "nones" have presented with equal difficulties and have not grasped the importance. This has been broached by looking at time series data of people and their religious affiliation over time, but only looked at how they move in and out of that group (Hout 2017; Lim et al. 2010). However, these studies do not look specifically at the measuring techniques of outside groups.

Importance of Measurement
From the very first attempts at measuring religious affiliation in the United States, the religiously unaffiliated have always seemed to be given short shrift. One of the first attempts to assess American religious demography was the United States Census in 1957, but it did not contain an option for those without a religious affiliation (Good 1959). This was also evident in other surveys from around the same time, as the secular respondents were lumped in together with other smaller religious groups (like Muslims and Buddhists), which creates a category that is completely non-sensical and has no real utility to social sci-

entists (Svalastoga 1965; Vernon 1968). This means that for nearly six decades the religiously non-affiliated were essentially ignored, and neglected as a group for analysis. This is evident from research that compared Gallup and GSS polls and found that the GSS had a larger share of "nones" than the Gallup poll primarily because it had a "no religion" option (Hout and Fischer 2002). This suggests that they have always existed, but were vastly undercounted by the most prominent surveys of the time.

It has also changed the way we qualitatively look at religious affiliation and "nones." McCaffree discusses the struggle between describing religiosity as a pattern of behavior or a belief system (2017). For example, social science has consistently found that people simply believe in a higher power but interpret the supernatural concepts through the lens of religious individualism (Bellah et al. 2007). It has also been described as an invisible religion, or non-doctrinal (Machalek and Martin 1976; Yinger 1969). Regardless, the religious "nones" represent a different type of religious belief and affiliation. This is noted by Schwadel who argues that having no religious affiliation changes the way people move through the world, and it can dramatically alter their political views and participation (Schwadel 2020).

Importance of Atheism Measurement
The above research and articles suggest that measuring the religious "nones" are important in understanding religious identity politics. What has not been properly discussed is how atheists fit into this equation, and how social science conceptualizes and measures their identification. Cragun began this work with his piece, where he delved into what groups are within the non-religious (2019). He asserts that many survey questions related to religion contain questions that are impossible for the non-religious to answer, and this practice keeps analysts from understanding the separate non-religious sects. Current works have neglected to look at how the aspects of religiosity, like belief, behavior, and belonging, work for non-religious people (Lee 2014; Schnell 2015). For example, deciding between believing or belonging, or double-barreled identification questions can force secular individuals into a religious category that they do not belong in (Converse 1986; Cragun 2016; Day 2011). This divide cuts both ways, however, as not all nonreligious people are atheists – for instance, agnostics clearly have a different belief structure and worldview (Kosmin et al. 2009; Lee 2014). Also, Cragun suggests that Atheists differ from other non-religious groups, through their belief in science, rejection of the supernatural, and criticism of other religions, and need to be studied apart from the group (Cragun 2014). Overall, Cragun suggests that the way the questions are formed in Gallup and other high-profile surveys force respondents into categories that in turn under measures their existence, both limiting our understanding of the group qualitatively and quantitatively (Cragun 2019).

In general, the study of what it means to chose an atheist affiliation (as opposed to agnostic or nothing in particular) is one that has been understudied by scholars of American religion. However, the Cooperative Congressional Election

Survey inadvertently offered researchers a unique insight into how atheists think about their place in the religious landscape of the United States when they did not include that option in the 2010 wave of their large, nationwide survey. If the option of "atheist" does not exist on the survey, one would assume that this group of people would choose an affiliation that is closer in proximity, such as "agnostic" or "nothing in particular." However, by reducing the number of "none" options from the three to two, that may nudge people away from choosing a religious non-affiliation and drive some potential atheists back into the Protestant or Catholic fold. Trying to understand the decision-making process that atheists navigated during this survey will be the focus of this research and will shine a light on how survey respondents engage with important questions regarding their religious affiliation.

Data

The Cooperative Congressional Election Study (CCES) began as a project out of Harvard University in 2006. The survey became popular quickly because of its federated style. For a fee, a team of researchers could add a battery of questions to the instrument which would be asked to one thousand respondents, while those one thousand respondents would also be asked a larger set of core questions related to basic demographics, political matters, as well as questions about various aspects of religiosity. Because of the open-source nature of the project, it became easy for dozens of research teams to sign on to the project, therefore the overall sample size began to far surpass most other surveys that are publicly available. For instance, the 2008 wave had 32,800 respondents. That jumped to 55,400 respondents in 2010, and 64,600 in 2016. The survey is conducted through an internet-based process which is facilitated by the polling firm YouGov using their pre-collected panels.

What is particularly helpful for students of American religion is that the CCES includes several questions related to all aspects of religiosity, but there is an especially robust battery of questions focused on religious belonging. The CCES adopted the Pew Research Center's approach to measuring religion, which begins with a broad question that asks, "What is your present religion, if any?" That is followed by twelve different choices: Protestant, Catholic, Mormon, Orthodox, Jewish, Muslim, Buddhist, Hindu, Atheist, Agnostic, Nothing in Particular, and Something Else. However, something curious happened in the 2010 wave of the CCES – the survey did not give respondents the option of choosing "atheist." It is simply missing from the data. While, on the surface, this looks like a mistake, it actually provides an interesting data puzzle for researchers: is it possible to reverse engineer the data to find the atheists in the 2010 wave?

What is the second choice of atheists?

A good starting point is to get a grasp on how the overall distribution of the sample shifted from the 2010 wave compared to the rest of the CCES samples from 2008 to 2018. To do that we calculated the share of the population that fell into each of the twelve religious categories in the seven waves in the CCES using the appropriate weights for each survey that were provided by the authors of the CCES. The 2010 wave is highlighted in **Figure 1** so it can be easily distinguished from the rest of the sample. In total, this data represents 378,156 total respondents to the CCES.

Using the rest of the CCES trend lines, what share of the 2010 wave of the CCES should we expect to have chosen the atheist option? While 3.4% of the population were atheists in 2008, that had jumped to 4.3% by 2012 and then steadily rose from that point to reach 6% by 2018. If we assume that the proportion of atheists in the 2010 wave was halfway between 2008 and 2012, then we can assume that about 3.8% of the population in the 2010 CCES were atheists, which represents about 2,100 total respondents.

It's also important to point out that there is an aberration in the 2008 data surrounding Protestant Christians. They were just 30.2% of the sample in 2008, which seems to be a dramatic outlier compared to the rest of the survey which pegs Protestants between 37–42% of the population. At the same time, the "something else" category was 21.3%, which is fifteen percentage points higher than the typical outcome. Because of these aberrations, it is not possible to detect whether atheists chose one of these two categories at a higher rate. Instead, we turn our attention to some of the most likely landing spots for potential atheists: agnostic or nothing in particular.

It appears that the trend for agnostics does see a bit of an outlier in 2010. While in 2008 they were 4.5% of the population, and that had climbed to 5% by 2012, then the 5.2% figure reported in 2010 seems to be slightly above the trendline. While it would appear that some atheists switched to the agnostic option, it was not that large. At the same time, the nothing in particular category does see a significant increase in 2010 compared to 2008 or 2012. While 14.4% of the population indicated that they were nothing in particular in 2008, that jumped 4.4 percentage points in 2010 to 18.8%, but then declined to 17.4% in the 2012 wave. In the rest of the series, there is no example of any religious groups increasing in size by four percentage points in one wave. If we can assume, consistent growth between waves, then the nothing in particular category would have been 15.9% in 2010. That 2.9 percentage point difference likely contains a bulk of the atheists who chose it as a backup option.

In terms of other groups, there was a possibly small increase in the share of Catholics. In 2010, there were 21% of the sample compared to 20.7% in 2008 and 19.1% in 2010. Although the overall portion of Catholics in the population stayed relatively stable from 2008 through 2016. The only other instance where there could be a possible increase is among people choosing the Jewish option. In 2010, 2.4% of respondents indicated that they were Jewish, which was a jump of about half a percentage point form 2008 and 2012. It seems possible that people who were ethnically Jewish but had atheist religious beliefs fell back to their ethnicity in 2010 when their religious preference was not an option.

Figure 1: Religious Demogrphy from the CCES (2010–2018).

Religious Importance

One place to narrow the search for miscategorized atheists would be to look at survey questions where this group would answer them in a distinct way from the rest of the sample. An ideal question is: "How important is religion in your life?" The possible response options range from "very important" to "not at all important." Atheists are a clear outlier when viewed through this lens with 95.3% of them indicating that religion is not important all, compared to 72.3% of agnostics, and just 16.4%

of the entire CCES from 2008–2018. It would be helpful to compare the distribution of people who chose "not important at all" in 2010 compared to the other waves. Religious categories that saw a significant rise would be likely places where atheists moved to in 2010. This is displayed in **Figure 2**.

In a typical year, between 22% and 27% of the sub-sample who say that religion is not important at all chose the atheist option. With that choice missing in the 2010 wave, there are significant shifts when it comes to the
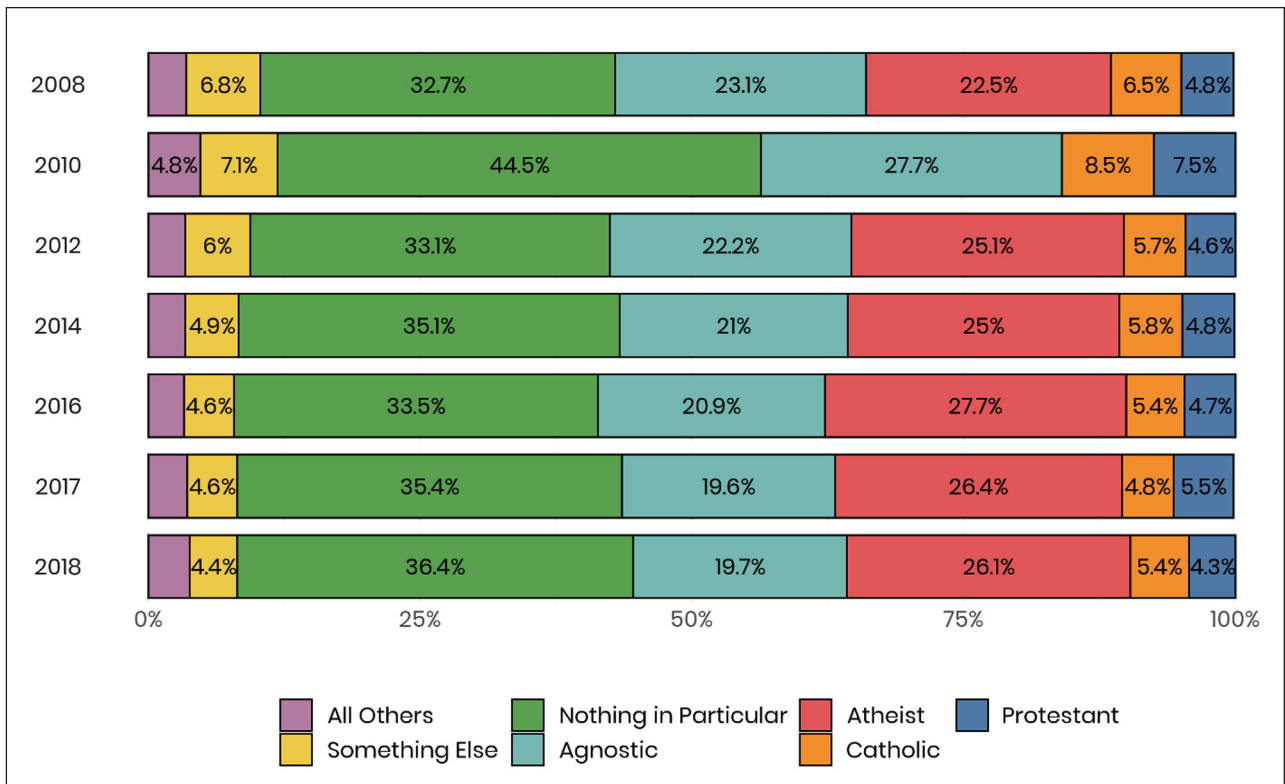
Figure 2: Religious Distribution of People Who Say Religion is Not Important.

distribution of other groups. For instance, the nothing in particular group is consistently around 33–34% of this subgroup, but jumped to 40.5% in 2010 – this six percent jump would likely be misplaced atheists. Other groups see smaller increases, for instance, agnostics rise just two to three percentage points compared to their typical share. That same increase of two to three points is also evident among Catholics. The rise in Protestants is somewhat larger at around four percentage points. From this view, it appears that the majority of atheists chose the nothing in particular option.

Church Attendance

One other good place to look for atheists in the 2010 sample is among people who never attend church services. In the entire CCES data, 88.9% of all atheists say that they never go to church compared to 69.5% of agnostics, 7.1% of Protestants and 10.4% of Catholics. As was done in the prior analysis, the samples for each wave were restricted to just people who never attended services and then the religious tradition was calculated for each of the years of the CCES.

The pattern in **Figure 3** is somewhat similar to the analysis that focused on just people who said that religion was not at all important to their lives. The nothing in particular category sees a significant boost in 2010 compared to the other years. Consistently, 35% of never attenders identified as nothing in particular, but that rose to 40.5% in 2010. Agnostics also see a small but noticeable increase – from a baseline of 14–15% to 17.4% in 2010. The other noticeable increase is among Protestants. In a typical year about 12% of people who never attend church identify as Protestants, but that increased to 16.3% in 2010 – that is

likely because of some atheists who chose the Protestant option.

Using Machine Learning to Find the Missing Atheists

However, it may be possible to use statistical techniques to reverse engineer the 2010 CCES sample to identify likely atheists. One of the most important innovations in the world of statistics and computing in recent years has been the adoption of machine learning methods. Their presence impacts nearly all aspects of life. From shopping websites suggesting potential add-on products before checkout, to social media websites recommending new people to friend or follow, artificial intelligence has become part of everyday life for most Americans. It can also be a potential solution to the problem of the misplaced atheists. Many of the most widely adopted algorithms in machine learning are focused on classification (Kotsiantis et al. 2007). For instance, a company wants to send a coupon to a consumer that is the most likely to use that enticement to make a purchase – identifying these potential customers out of databases containing millions of data points by hand would be impractical. However, machine learning can quickly iterate over thousands of possible variables and arrive at a solution that can be constantly refined based on feedback generated by consumer behavior.

Social science has just begun to adopt machine learning techniques to help with classification problems. Clustering techniques have been employed on large survey datasets as a means of sorting respondents into different religious groups (Pearce and Denton 2011; Storm 2009). Researchers have begun to use machine learning as
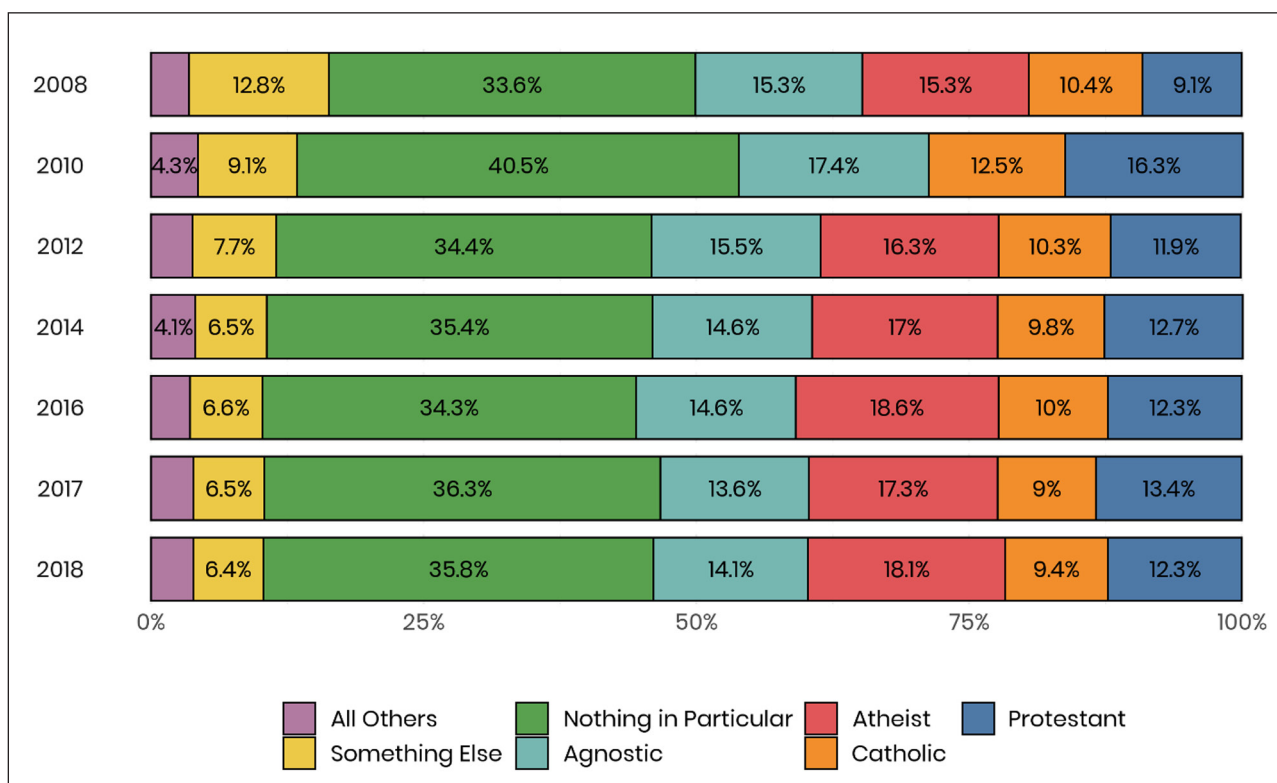
**Figure 3:** Religious Distribution of Those Who Never Attend Church.

a means to generate coding for content analysis on a large scale (Scharkow 2013; Burscher 2015). Recently, a team of political scientists found that random forests were more effective at identifying the onset of civil war than a traditional logistic regression (Muchlinski et al. 2016). Some of the most prominent methodologists in the social sciences have begun to develop frameworks and best practices for integrating machine learning into traditional academic research (Grimmer 2015).

The problem that a researcher is confronted with in the 2010 CCES can be approached by using a machine learning technique called random forests. A random forest classifier is based on a simple machine learning principle – a decision tree. A decision tree begins by finding a variable in the data that will divide the sample in the most distinct way possible. The creation of these decision trees occurs millions of times in a random forest model, trying to find a series of bifurcations where the trees correctly predict the outcome 100% of the time (Liaw and Wiener 2002). To accomplish this, a dataset is divided up into a training dataset and a test dataset (Ham et al. 2005). The training dataset has the outcome already labeled. In this case, a dichotomous variable was created which separated atheists from all other religious traditions. This training data was the 2008 and 2012 CCES waves, which included the atheist option in the religious tradition question. The labeled data from 2008 and 2012 also included several variables: church attendance, importance of religion, frequency of prayer, born-again status, partisanship, ideology, age, race, gender, education, marital status, and income. These variables were used in a set of 100 decision trees to train the algorithm to make correct guesses as frequently as possible. One of the benefits of random

forests is that the algorithm determines which variables are the most important to generating correct guesses and excludes those factors that do not increase the model's accuracy (Archer and Kimes 2008). The algorithm was able to construct a decision tree that correctly classified atheists in the training data 96.3% of the time.

One of the outputs from a random forest model is a ranking of how important each variable was in terms of its ability to generate correct predictions. The variables that the model relied on the most were frequency of prayer, religious importance, age, income, and church attendance. In general, demographic variables were less helpful to the model, while variables that related to religiosity were particularly valuable. Once the random forest had been developed using the training data from the 2008 and 2012 CCES, it was then used to predict whether the 55,400 respondents in the 2010 CCES test data were atheists or not. The result is a score for each respondent ranging from zero (meaning the model predicts that there is no chance that the person is an atheist) to one (which indicates a high probability that the individual identifies as an atheist).

Recall that the expected share of the population that should have chosen an atheist affiliation in 2010 was approximately 3.8%. Therefore, the sample was restricted to those that the algorithm scored with the highest likelihood of being an atheist until 3.8% of the total population was selected. In general, the random forest model was effective at creating a sample of predicted atheists in 2010 that looked like those who chose the atheist option in 2008 or 2012. As can be seen in **Figure 4**, for most of the key variables, there was very little difference between the two groups. In fact, the only variable where the
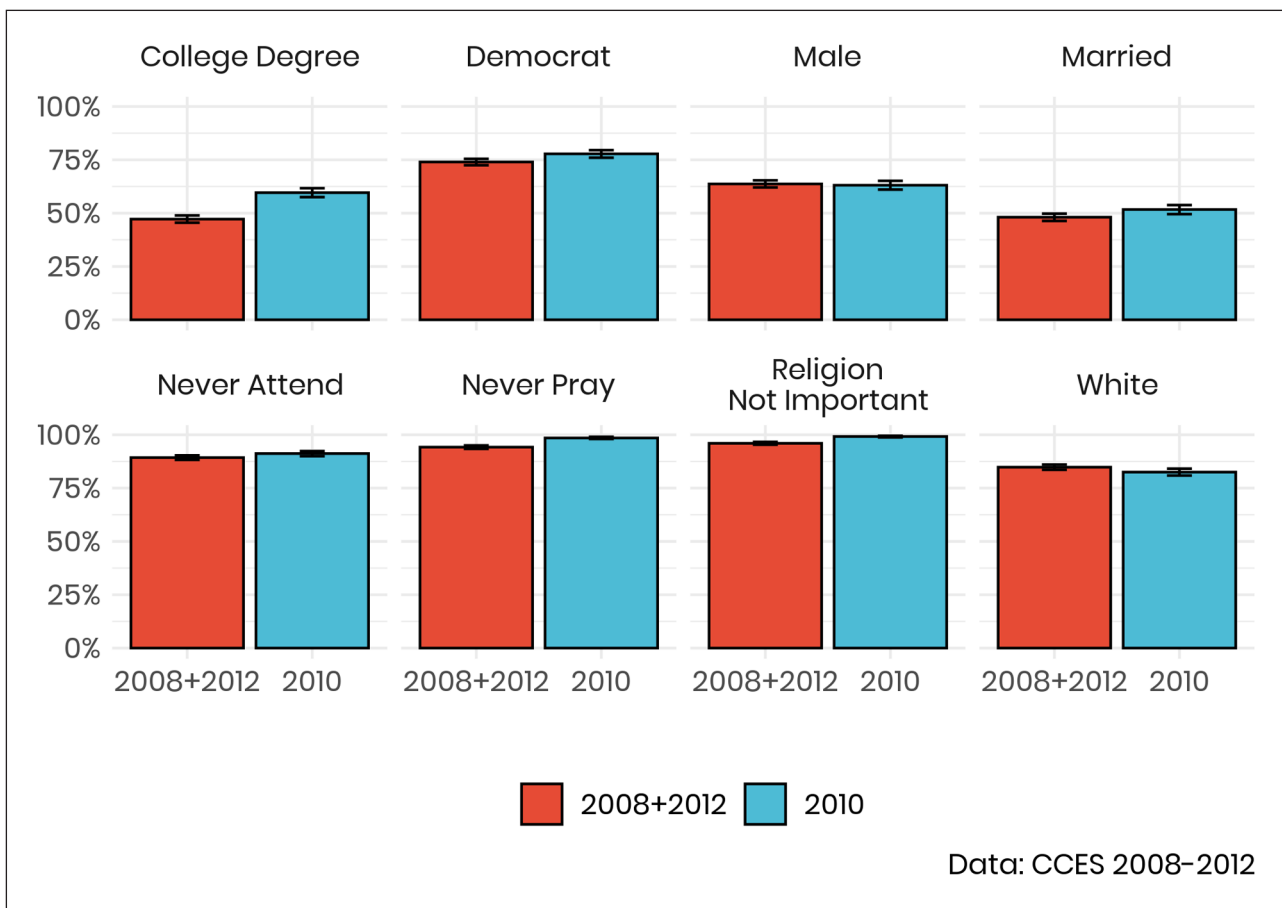
Figure 4: Predicted Atheists in 2010 vs. Atheist in 2008/2012.

divergence was substantively large was education. In the predicted sample of atheists, 59.6% had a college degree compared to 47.2% of those in 2008 and 2012.

Using those who scored in the highest 3.8% of likelihood to be an atheist using the random forest algorithm, it becomes possible to determine how this share of the population navigated the religious tradition question in 2010. The data tells us a clear story, which is visualized in **Figure 5**: the vast majority of potential atheists chose another type of religiously unaffiliated tradition. Two out of four chose the agnostic option, while the same share picked nothing in particular. Therefore, the total share of nones was only slightly diminished due to this survey error. It is worth noting that even though agnostics seemed like the most attractive landing spot for many of these atheists, the data indicates that they were just as likely to pick either of the religiously unaffiliated options. It is also worth pointing out that very few of these predicted atheists chose Judaism or Christianity. Recall that there was a slight bump in the share of Jews, Catholics and Protestants in 2010, and that can be somewhat attributed to the fact that between 3.3% and 5% of potential atheists chose one of these options.

Conclusion

Looked at from a strictly empirical spectrum, there is a clear continuum of the religiously unaffiliated. Nine in ten (consistency, numbers or words, so far you used numbers) atheists never attend church services, compared to seven in ten agnostics, and half of nothing in particulars. Ninety-five percent of atheists say that religion is not at all important in their lives, while 72.3% of agnostics, and just 36.1% of nothing in particulars respond in the same manner. On dimensions of political partisanship and public opinion, the same order emerges – atheists are further to the left, followed by agnostics, and nothing in particulars are more in the center of the political spectrum (Schwadel 2020). One could assume from this that if atheists were left to pick their second choice of religious affiliation, then agnostics would be the clear choice. However, that is not the conclusion from the data. While 80% of misplaced atheists still stayed inside the religiously unaffiliated grouping, they were evenly split between the agnostic and nothing in particular option. It would appear that at the ground level, atheists do not see the continuum in the same way that the data does. As such, it does not appear that social science can assume that atheists see agnostics as their closely related cousins. It is possible that some atheists actively reject an agnostic worldview and that led to these results.

This is something worth some careful thought for religious demographers when they consider assessing the size and composition of the religious "nones" in the United States. The three categories of atheist, agnostic, and nothing in particular have become the standard way to classify the religiously unaffiliated. But, is that the most accurate way of conceptualizing nonreligion? The survey question seems to be straightforward: what is your present religion,
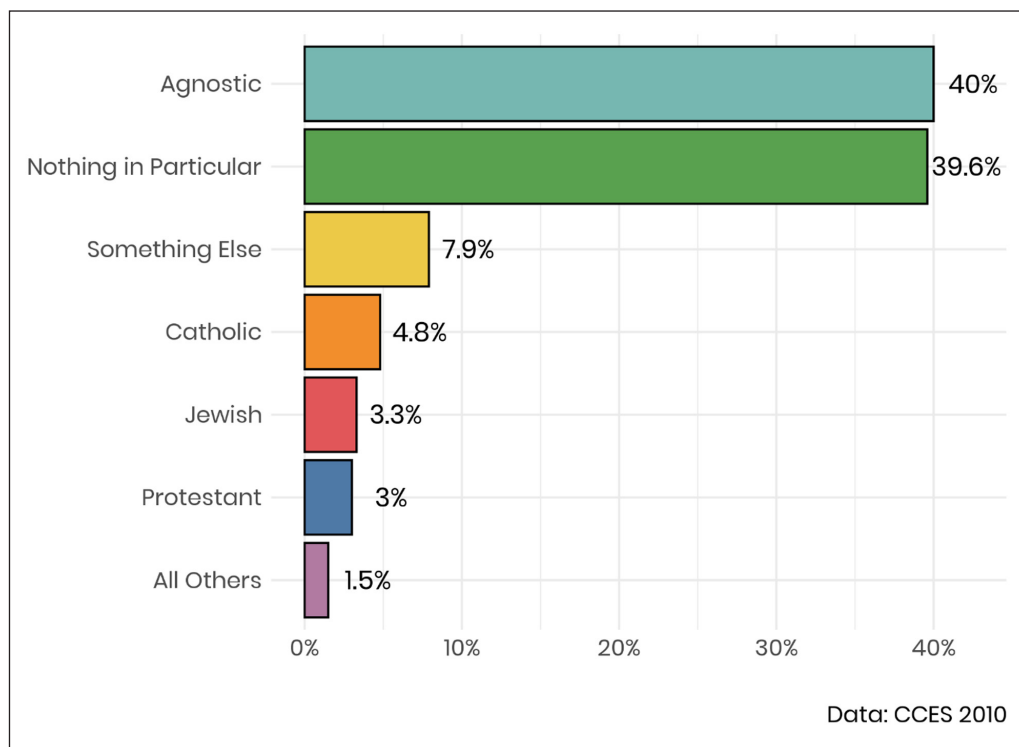
Figure 5: What Did Atheists Choose in 2010?

if any? However, the response options run the gamut from religious affiliation (Protestant) to an ethnic group (Jewish) to a belief orientation (atheism, agnosticism). Identifying as a Catholic can mean a wide variety of things to a respondent, such as regular attendance at Mass, or the desire to have a priest administer last rites if they were on their deathbed. Or it could mean that their family has Irish or Italian origins, which are often deeply intertwined with the Catholic church. The label of "atheist" carries with it an entirely different set of implications. At its core, the term refers to a belief system, not necessarily a societal group that has culture and norms. That is not a small difference. However, this entire discussion returns back to a central, pressing question for those who study the American religious landscape: how does the average person understand American religion and is that at odds with how researchers conceptualize and measure it? More work is needed to validate our current survey measures.

In addition to the contribution of this work to the understanding of religious disaffiliation in the United States, we hope that this will give some social scientists a gentle introduction to the world of machine learning and artificial intelligence. There are specific problems that social science faces which could be addressed, at least in part, by the application of algorithms. However, it is crucial to note that machine learning is not some type of panacea for the social sciences. It is crucial to apply machine learning techniques when they are the most appropriate. For instance, they are not well suited to testing questions of causal inference because many of them tend to have a "black box" quality, whereby an analyst can see the result of the machine learning algorithm, but the model does not fully explain how it arrived at that conclusion (Rudin 2019). However, these algorithms can help to turn our

attention to variables or possible groupings of individuals that have not been considered by social science before. Scholars in social science are uniquely equipped to use the results of these machine learning techniques as opportunities to generate potential hypotheses that can be tested with more traditional techniques, such as regression. As such, we can illuminate more of the social world and gain a richer understanding of religiosity.

Competing Interests

References
Archer, KJ and **Kimes, RV.** 2008. Empirical Characterization of Random Forest Variable Importance Measures. *Computational Statistics & Data Analysis*, 52(4): 2249–2260. DOI: https://doi.org/10.1016/j.csda.2007.08.015

Baker, J and **Smith, B.** 2015. *American Secularism: Cultural Contours of Nonreligious Belief Systems*. NYU Press.

Bellah, RN, et al. 2007. *Habits of the Heart, With a New Preface: Individualism and Commitment in American Life*. Berkeley, Calif. Los Angeles, Calif. London: University of California Press.

Burscher, B, Vliegenthart, R and **De Vreese, CH** 2015. Using Supervised Machine Learning To Code Policy Issues: Can Classifiers Generalize Across Contexts? *The ANNALS of the American Academy of Political and Social Science*, 659(1): 122–131. DOI: https://doi.org/10.1177/0002716215569441

Caplow, T. 1998. The Case of the Phantom Episcopalians. *American Sociological Review*, 63(1): 112–113. DOI: https://doi.org/10.2307/2657481

Converse, JM. 1986. *Survey Questions: Handcrafting the Standardized Questionnaire.* Beverly Hills: Sage Publications.

Cragun, RT. 2014. Who Are the 'New Atheists'? In: *Atheist Identities: Spaces and Social Contexts,* Beamon, L and Tomlins, S (eds.), 195–211. New York: Springer. DOI: https://doi.org/10.1007/978-3-319-09602-5_12

Cragun, RT. 2016. Defining That Which is "Other to" Religion: Secularism, Humanism, Atheism, Freethought, etc. In: Zuckerman, P (ed.), *Religion: Beyond Religion,* 1–16. New York City: MacMillan Publishing Company (Macmillan Reference USA).

Cragun, RT. 2019. Questions You Should Never Ask an Atheist: Towards Better Measures of Nonreligion and Secularity. *Secularism and Nonreligion,* 8(6). DOI: https://doi.org/10.5334/snr.122

Day, A. 2011. *Believing in Belonging: Belief and Social Identity in the Modern World.* Oxford; New York: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199577873.001.0001

Good, D. 1959. Questions on Religion in the United States Census. *Population Index,* 25(1): 3–16. DOI: https://doi.org/10.2307/2731545

Grimmer, J. 2015. We Are All Social Scientists Now: How Big Data, Machine Learning, and Causal Inference Work Together. *PS: Political Science & Politics,* 48(1): 80–83. DOI: https://doi.org/10.1017/S1049096514001784

Hadaway, C, Kirk, P and **Chaves, M.** 1993. What the Polls Don't Show: A Closer Look at US Church Attendance. *American Sociological Review,* 58(6): 741–752. DOI: https://doi.org/10.2307/2095948

Hadaway, CK, Marler, PL and **Chaves, M.** 1998. Overreporting Church Attendance in America: Evidence That Demands the Same Verdict. *American Sociological Review,* 63(1): 122–130. DOI: https://doi.org/10.2307/2657484

Hall, DE, Koenig, HG and **Meador, KG.** 2009. Hitting the Target: Why Existing Measures of "Religiousness" are Really Reverse-Scored Measures of "Secularism." *Explore (New York, N.Y.),* 4(6): 368–73. DOI: https://doi.org/10.1016/j.explore.2008.08.002

Hall, DE, Meador, KG and **Koenig, HG.** 2008. 'Measuring Religiousness in Health Research: Review and Critique' *Journal of Religion and Health,* 47(2): 134–163. DOI: https://doi.org/10.1007/s10943-008-9165-2

Ham, J, Chen, Y, Crawford, MM and **Ghosh, J.** 2005. Investigation of the Random Forest Framework for Classification of Hyperspectral Data. *IEEE Transactions on Geoscience and Remote Sensing,* 43(3): 492–501. DOI: https://doi.org/10.1109/TGRS.2004.842481

Hout, M. 2017. Religious Ambivalence, Liminality, and the Increase of No Religious Preference in the United States, 2006–2014. *Journal for the Scientific Study of Religion,* 56(1): 52–63. DOI: https://doi.org/10.1111/jssr.12314

Hout, M and **Fischer, CS.** 2002. Why More Americans Have No Religious Preference: Politics and Generations. *American Sociological Review,* 67(2): 165. DOI: https://doi.org/10.2307/3088891

Hout, M and **Greeley, A.** 1987. The Center Doesn't Hold: Church Attendance in the United States, 1940–1984. *American Sociological Review,* 325–345. DOI: https://doi.org/10.2307/2095353

Hout, M and **Greeley, A.** 1998. What Church Officials' Reports Don't Show: Another Look at Church Attendance Data. *American Sociological Review,* 63(1): 113–119. DOI: https://doi.org/10.2307/2657482

Hwang, K, Hammer, JH and **Cragun, RT.** 2011. Extending Religion-Health Research to Nontheistic Minorities: Issues and Concerns. *Journal of Religion and Health,* 50(3): 608–22. DOI: https://doi.org/10.1007/s10943-009-9296-0

Karp, JA and **Brockington, D.** 2005. Social Desirability and Response Validity: A Comparative Analysis of Overreporting Voter Turnout in Five Countries. *The Journal of Politics,* 67(3): 825–840. DOI: https://doi.org/10.1111/j.1468-2508.2005.00341.x

Kohut, A, Green, JC, Keeter, S and **Toth, RC.** 2001. *The Diminishing Divide: Religion's Changing Role in American Politics.* Washington: Brookings Institution Press.

Kosmin, BA, et al. 2009. *American Nones: The Profile of the No Religion Population.* Hartford, CT: Institute for the Study of Secularism in Society and Culture, 29.

Kotsiantis, SB, Zaharakis, I and **Pintelas, P.** 2007. Supervised Machine Learning: A Review of Classification Techniques. *Emerging Artificial Intelligence Applications in Computer Engineering,* 160: 3–24. DOI: https://doi.org/10.1007/s10462-007-9052-3

Lee, L. 2014. Secular or nonreligious? Investigating and Interpreting Generic "Not Religious" Categories and Populations. *Religion,* 44(3): 466–482. DOI: https://doi.org/10.1080/0048721X.2014.904035

Liaw, A and **Wiener, M.** 2002. Classification and Regression by randomForest. *R News,* 2(3):18–22.

Lim, C, MacGregor, CA and **Putnam, RD.** 2010. Secular and Liminal: Discovering Heterogeneity Among Religious Nones. *Journal for the Scientific Study of Religion,* 49(4): 596–618. DOI: https://doi.org/10.1111/j.1468-5906.2010.01533.x

Machalek, R and **Martin, M.** 1976. 'Invisible' Religions: Some Preliminary Evidence. *Journal for the Scientific Study of Religion,* 15(4): 311–321. DOI: https://doi.org/10.2307/1385634

McCaffree, K. 2017. *The Secular Landscape: The Decline of Religion in America.* Berlin: Springer.

McGuire, WJ. 1968. Personality and Attitude Change: An Information-Processing Theory. *Psychological Foundations of Attitudes,* 171: 196. DOI: https://doi.org/10.1016/B978-1-4832-3071-9.50013-1

Muchlinski, D, Siroky, D, He, J and **Kocher, M.** 2016. Comparing Random Forest with Logistic Regression for Predicting Class-Imbalanced Civil War Onset Data. *Political Analysis,* 24(1), 87–103. DOI: https://doi.org/10.1093/pan/mpv024

Nederhof, AJ. 1985. Methods of Coping with Social Desirability Bias: A Review. *European Journal of Social Psychology,* 15(3): 263–280. DOI: https://doi.org/10.1002/ejsp.2420150303

Nederhof, AJ and **Zwier, AG.** 1983. The 'Crisis' in Social Psychology, an Empirical Approach. *European Journal of Social Psychology*, 13(3): 255–280. DOI: https://doi.org/10.1002/ejsp.2420130305

Pearce, L and **Denton, ML.** 2011. *A Faith of Their Own: Stability and Change in the Religiosity of America's Adolescents.* Oxford University Press.

Presser, S and **Stinson, L.** 1998. Data Collection Mode and Social Desirability Bias in Self-Reported Religious Attendance. *American Sociological Review*, 63(1): 137–145. DOI: https://doi.org/10.2307/2657486

Rudin, C. 2019. Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. *Nature Machine Intelligence*, 1(5): 206–215. DOI: https://doi.org/10.1038/s42256-019-0048-x

Scharkow, M. 2013. Thematic Content Analysis Using Supervised Machine Learning: An Empirical Evaluation Using German Online News. *Quality & Quantity*, 47(2): 761–773. DOI: https://doi.org/10.1007/s11135-011-9545-7

Schnell, T. 2015. Dimensions of Secularity (DoS): An Open Inventory to Measure Facets of Secular Identities. *The International Journal for the Psychology of Religion*, 25(4): 272–292. DOI: https://doi.org/10.1080/10508619.2014.967541

Schwadel, P. 2020. The Politics of Religious Nones. *Journal for the Scientific Study of Religion*, 59(1). DOI: https://doi.org/10.1111/jssr.12640

Smith, TW. 1991. Counting Flocks and Lost Sheep: Trends in Religious Preference Since World War II. *GSS Social Change Report*, 26: 40–42.

Smith, TW and **Seokho, K.** 2007. C*ounting Religious Nones and Other Religious Measurement Issues: A Comparison of the Baylor Religion Survey and General Social Survey.* GSS Methodological Report No.110:17.

Smith, TW and **Seokho, K.** 2005. The Vanishing Protestant Majority. *Journal for the Scientific Study of Religion*, 44(2): 211–223. DOI: https://doi.org/10.1111/j.1468-5906.2005.00277.x

Storm, I. 2009. Halfway to Heaven: Four Types of Fuzzy Fidelity in Europe. *Journal for the Scientific Study of Religion*, 48(4): 702–18. DOI: https://doi.org/10.1111/j.1468-5906.2009.01474.x

Streb, MJ, Burrell, B, Frederick, B and **Genovese, MA.** 2007. Social Desirability Efects and Support for a Female American President. *Public Opinion Quarterly*, 72(1): 76–89. DOI: https://doi.org/10.1093/poq/nfm035

Svalastoga, K. 1965. *Social Differentiation.* New York: D. McKay Co.

Swatos, WH and **Christiano, KJ.** 1999. Secularization Theory: The Course of a Concept. *Sociology of Religion*, 60(3): 209–228. DOI: https://doi.org/10.2307/3711934

Vernon, GM. 1968. The Religious "Nones": A Neglected Category. *Journal for the Scientific Study of Religion*, 7(2): 219–229. DOI: https://doi.org/10.2307/1384629

Woodberry, RD. 1998. When Surveys Lie and People Tell the Truth: How Surveys Oversample Church Attenders. *American Sociological Review*, 63(1): 119–122. DOI: https://doi.org/10.2307/2657483

Yinger, JM. 1969. A Structural Examination of Religion. *Journal for the Scientific Study of Religion*, 8(1): 88–99. DOI: https://doi.org/10.2307/1385257

Zuckerman, P, Galen, LW and **Pasquale, FL.** 2016. *The Nonreligious: Understanding Secular People and Societies.* New York: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199924950.001.0001